Schuyler D. Smith, Ph.D.

✉ schuyler.d.smith@gmail.com
 github.com/schuyler-smith
 Google Scholar
 schuyler-smith.github.io
 linkedin.com/in/schuyler-ds
📞 +1-(413)-212-9110
 Durham, NC 27713

Education

Ph.D.   **Bioinformatics and Computational Biology**   December 2021
Iowa State University, Ames, IA
Characterization of environmental microbiomes impacted by Iowa agriculture.

M.S.   **Quantitative Genetics and Plant Breeding**   May 2015
Texas A&M University, College Station, TX
Molecular Characterization of the Texas Maize Breeding Program.

B.S.   **Genetics**   December 2012
Iowa State University, Ames, IA

Experience

Senior Computational Scientist, Bioinformatics   January 2022 – *present*
Nutrien Ag Solutions (Waypoint Analytical) | Soil Biome | Champaign, Illinois
- Developmed end-to-end NGS data pipelines for metagenomic samples.
- Built and maintained NextFlow workflows in cloud-computing environments on AWS and GCP.
- Developed custom software in C⁺ for lab processing of qPCR data.
- Developed packages in both R and python for advanced microbiome data analyses.
- Collect and Maintained databases for both lab and computational teams.
- Constructed soil-biome report generation platform.
- Implemented models and algorithms for product-placement using microbiome metrics.
- Implemented RAG and GenAI systems for custom chatbot apps.
- Implemented gRPC server and protocols for API data system for biome team.
- Project lead for development of metagenomic results discovery database.
- Developed analyses and presentations for stakeholders.
- Mentored junior colleagues and summer interns.

Graduate Research Assistant - Ph.D.   January 2017 – December 2021
Iowa State University | GERMS Laboratory | Ames, Iowa
- Developed pipeline for processing 16S sequencing.
- Authored R-packages for microbiome data analysis.
- Published research on antibiotic-resistance gene dispersal in environmental systems.
- Conducted research on harmful algal bloom biome-community and predictors.
- Published research evaluating large-scale marker detection for microbiome genes.
- Mentored junior lab members.
- Instructor for Data & Software Carpentries workshops.

Graduate Research Assistant - Ph.D.   June 2015 – September 2016
University of Wisconsin - Madison | Potato Genetics Laboratory | Madison, Wisconsin
- Developed pipeline for NGS data and SNP-calling for autotetraploid crops.
- Authored R-packages for ML methods for autotetraploid genotyping.

Biological Science Technician - Internship   June 2014 - December 2014
USDA-ARS | Arid-Land Agricultural Research Cente | Maricopa, Arizona
- Evaluated use of image-based high-throughput phenotyping platforms.

Graduate Research Assistant - M.S.   January 2013 - May 2015
Texas A&M University | Maize Genetics Laboratory | College Station, Texas
- Published research characterizing markers of Texas maize germplasm.

Maize Breeding Intern   May 2012 - September 2012
  Monsanto Company | Huxley Research Station | Huxley, Iowa
Maize Product Trait Development Intern – 6 month   June 2011 - December 2011
  DuPont Pioneer | Willmar Research Station | Willmar, Minnesota

Applicable Skills

python	AWS / GCP / SLURM	Shiny / Dash / Power BI
R	NextFlow / SnakeMake / AirFlow	Markdown / HTML / CSS
C++	CI / CD (Git / Github Actions)	LaTeX
Shell	Docker / Singularity	Image J
SQL	gRPC	Linux / Windows / MacOS

Software

phylosmith: R package for reproducible and efficient microbiome analysis with phyloseq-objects.

smartchip_analyzer: C++ program for processing data from a SmartChip qPCR.

weather_api: python package with modules for accessing a weather-data API.

ssBLAST: R package wrapping functions written in C++ to parse BLAST outputs and manipulate corresponding FASTQ/FASTA files.

simple.dada: R package for streamlined implementation of the dada2 processing pipeline.

schuylR: R package of various functions that I often find useful.

schemeR: R package for my ggplot2 graph theme.

Workflows: Workflows I am building for various bioinformatics analyses.

Publications

Smith, **S. D.**, Velásquez-Zapata, V., & Wise, R. P. (2024). *Ngpint_v3: A containerized orchestration software for discovery of next-generation protein–protein interactions*. *SSRN 4955985*.

Velásquez-Zapata, V., Elmore, J. M., Patel, S., **Smith**, **S. D.**, Fuerst, G., & Wise, R. P. (2024). *Effector-host protein networks in an integrated barley-powdery mildew interactome*. *Plant and Animal Genome Conference/PAG 31 (January 12-17, 2024)*.

Velásquez-Zapata, V., **Smith**, **S. D.**, Surana, P., Chapman, A. V., Jaiswal, N., Helm, M., & Wise, R. P. (2024). *Diverse epistatic effects in barley-powdery mildew interactions localize to host chromosome hotspots*. *Iscience*, *27*(10).

Velásquez-Zapata, V., **Smith**, **S. D.**, Surana, P., Chapman, A., & Wise, R. (2023). *Transcriptome-based host epistasis and pathogen co-expression in barley-powdery mildew interactions* [Preprint]. *bioRxiv*.

Smith, **S. D.**, Choi, J., Ricker, N., Yang, F., Hinsa-Leasure, S., Soupir, M. L., Allen, H. K., & Howe, A. (2022). *Diversity of antibiotic resistance genes and transfer elements-quantitative monitoring (darte-qm): A method for detection of antimicrobial resistance in environmental samples*. *Communications Biology*, *5*(1), 216.

Yu, W., Lawrence, N., Sooksa-Nguan, T., **Smith**, **S. D.**, Tenesaca, C., Howe, A., & Hall, S. (2021). *Microbial linkages to soil biogeochemical processes in a poorly drained agricultural ecosystem*. *Soil Biology and Biochemistry*, *156*, 108228.

Smith, **S. D.** (2019). *Phylosmith: An r-package for reproducible and efficient microbiome analysis with phyloseq-objects*. *Journal of Open Source Software*.

Smith, **S. D.**, Colgan, P., Yang, F., Rieke, E., Soupir, M., Moorman, T., Allen, H., & Howe, A. (2019). *Investigating the dispersal of antibiotic resistance associated genes from manure application to soil and drainage waters in simulated agricultural farmland systems*. *PLoS One*, *14*(9), e0222470.

Choi, J., Rieke, E. L., Moorman, T. B., Soupir, M. L., Allen, H. K., **Smith**, **S. D.**, & Howe, A. (2018). *Practical implications of erythromycin resistance gene diversity on surveillance and monitoring of resistance*. *FEMS microbiology ecology*, *94*(4), fiy006.

Smith, **S. D.**, Heffner, E., & Murray, S. C. (2015). *Molecular analysis of genetic diversity in a texas maize breeding program*. *CORE*.

Workshops Taught

Data Carpentry: Genomics | Oklahoma State University | online	2021
Software Carpentry: Shell, Git, R for Reproducible Scientific Analysis | University of Idaho | online	2021
Data Carpentry: Ecology | Merck & Co. | online	2020
Data Carpentry: Genomics | George Washington University | Washington, D.C.	2020
Data Carpentry: Ecology | Merck & Co. | Upper Gwynedd, PA	2019
Girls in Science Initiaive | Science Center of Iowa | Des Moines, IA	2019
Introduction to Data Analysis | Iowa State University | Ames, IA	2017 & 2018
Iowa State BCB Data Analysis Language Workshops Introduction to Unix	2017 & 2018
Iowa State BCB Data Analysis Language Workshops Introduction to Python	2017
Explorations in Data Analyses for Metagenomic Advances in Microbial Ecology	2017 & 2018

Graduate Coursework Completed

Ph.D. Bioinformatics and Computational Biology:
- Bioinformatic Algorithms
- Statistical Bioinformatics
- Bioinformatic Systems
- Genomic Sciences
- Linear Mixed Models
- Fundamentals of Predictive Plant Phenomics
- Tools for Reproducible Research
- Plant Genetics
- Biometric Procedures in Plant Breeding
- Advanced Plant Breeding
- Selection Theory

M.S. Plant Breeding and Quantitative Genetics:
- Quantitative Genetics
- Statistics in Research I
- Statistics in Research II
- Plant Breeding I
- Plant Breeding II
- Experimental Design
- Molecular and Quantitative Genetics in Plant Breeding
- Host-Plant Resistance